Mathematical Statistics **II**
Spring 2009
Solutions to the in-class assignments (04/06/09).

1. Suppose that X_1, X_2, X_3 and X_4 are independent and identically distributed from a normal distribution with mean μ and variance 9. Then we know that the maximum likelihood estimator of μ, $\overline{X}$, has normal distribution with mean μ and variance 9/4. Let the random variable Z be defined as
$$Z = \frac{\overline{X} - \mu}{\sqrt{9/4}}.$$
Then $Z \sim N(0,1)$ and it follows that
$$P\left(\overline{X} - \tfrac{3}{2}(1.96) \leq \mu \leq \overline{X} + \tfrac{3}{2}(1.96)\right) = 0.95$$
since $\Phi^{-1}(0.025) = -1.96$ where $\Phi(\cdot)$ is the CDF of $N(0,1)$. So a 95% confidence interval for μ is $\left(\overline{X} - \frac{3}{2}(1.96),\ \overline{X} + \frac{3}{2}(1.96)\right)$. By taking different values of μ, say $\mu = 3, 10, 15, 20$, verify numerically that the coverage probability of the above confidence interval is 95%.

Solution. Using the following R code:

```
muvalues<-c(3,10,15,20); # Differnt values for mu
SimSize<-1000 # Simulation Size
n<-4
sigma<-3  #Standard Deviation
Result<-c() # Storage for the results for each value of mu
# Looping Thru mu starts here
for (mu in c(3,10,15,20))
{
R<-c()
for (i in 1:SimSize)
{X<-rnorm(n,mu,sigma)
xbar<-mean(X)
LL<-xbar-(sigma/sqrt(n))*1.96
UL<-xbar+(sigma/sqrt(n))*1.96
R<-cbind(R, (LL < mu) && (UL > mu))
}
Result<-rbind(Result,c(mu,mean(R)*100))
}
Result<-as.data.frame(Result) # Converts matrix into data frame.
names(Result)<-c("mu","CP(%)") #Changes the names of the variables
Result
```

when run once resulted

```
  mu CP(%)
1  3  95.0
2 10  94.8
3 15  95.1
4 20  95.7
```

$\square$

2. Let $X_1, X_2, \ldots, X_n$ be a sample from $N(\mu, \sigma^2)$ where both μ and σ^2 are unknown. Suppose we are interested in constructing a $100\gamma\%$a shortest-length confidence interval for μ. Let us define the random variable T as

$$T = \frac{(\overline{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{\frac{(n-1)^2}{\sigma^2} S^2/(n-1)}} = \frac{\sqrt{n}(\overline{X} - \mu)}{S} \sim t_{n-1}.$$

Let q_1 and q_2 be two numbers such that

$$P\left(q_1 < \frac{\overline{X} - \mu}{S}\sqrt{n} < q_2\right) = P\left(\overline{X} - q_2 \frac{S}{\sqrt{n}} < \mu < \overline{X} - q_1 \frac{S}{\sqrt{n}}\right) = \gamma.$$

We wish to minimize the expected length, $E(L)$,

$$E(L) = (q_2 - q_1)\frac{E(S)}{\sqrt{n}}$$

subject to

$$\int_{q_1}^{q_2} f_T(t)\, dt = \gamma,$$

where f_T is the probability density function of T. Notice that the distribution of T is symmetric about zero. So the values q_1 and q_2 which minimize $q_2 - q_1$ subject to area under the t_{n-1} curve between q_1 and q_2 being γ must satisfy $q_1 = -q_2$. Therefore,

$$q_1 = -q_2 = t_{\frac{1-\gamma}{2}, n-1}$$

where $t_{\alpha,n-1}$ is the αth quantile of the t distribution with $n-1$ degrees of freedom.